The Number of Hierarchical Orderings

N. J. A. Sloane
AT&T Shannon Labs
Florham Park, NJ 07932–0971, USA
(njas@research.att.com, http://www.research.att.com/∼njas)

Thomas Wieder
Darmstadt University of Technology
Institute for Materials Science
D–64287 Darmstadt, Germany
(thomas.wieder@epost.de, http://homepages.tu-darmstadt.de/∼wieder)

Abstract

An ordered set-partition (or preferential arrangement) of n labeled elements represents a single "hierarchy"; these are enumerated by the ordered Bell numbers. In this note we determine the number of "hierarchical orderings" or "societies", where the n elements are first partitioned into $m \le n$ subsets and a hierarchy is specified for each subset. We also consider the unlabeled case, where the ordered Bell numbers are replaced by the composition numbers. If there is only a single hierarchy, we show that the average rank of an element is asymptotic to $n/(4 \log 2)$ in the labeled case and to $n/4$ in the unlabeled case.

Keywords: ordered set-partition, preferential arrangement, hierarchical ordering, society

AMS 2000 Classification: Primary 06A07

1 Introduction

Suppose we are given a set S of n labeled elements (or "individuals"). The number of ordered set-partitions (or "hierarchies") on S is given by the ordered Bell number B_n (see sequence A670 in [10] for properties and references). A *hierarchical ordering* or *society* on S is specified by first distributing the elements into $m \le n$ unlabeled and nonempty subsets, and forming an ordered set-partition in each subset. In Section 2 we will determine the number of different structures of this type. Section 3 discusses the analogous question in the unlabeled case, when the elements are indistinguishable.

The original motivation for this work was to try to describe the structure of a "typical" society. Of course Sections 2 and 3 just enumerate them. However, in Section 4 we consider the distribution of ranks in a random selection of a single hierarchy, in both the labeled and unlabeled cases.

Structures of the kind considered here were discussed in a classic combinatorics paper by Motzkin [6]. In his terminology, the labeled structures in Section 2 would be called "sets of lists of sets", and the unlabeled structures in Section 3 "sets of lists of numbers".*

*The Maple package *combstruct* [5] makes it easy to generate such structures. For example, the labeled

2 The labeled case

Let H_n denote the number of possible hierarchical orderings or societies, with exponential generating function (or e.g.f.) $H(x) = \sum_{n \ge 0} H_n x^n / n!$. An explicit formula for $H(x)$ follows from a standard application of the exponential formula in combinatorics.

Theorem 1

$$H(x) \;=\; \exp(B(x) - 1)\ , \tag{1}$$

where $B(x) = 1/(2 - e^x)$ *is the e.g.f. for the ordered Bell numbers.*

Proof. This is immediate from (for example) [11, Cor. 3.4.1], or [2, Chap. 1.4, p. 46]. ∎

The first few values H_n for $n = 0, 1, 2, \ldots$ are $1, 1, 4, 23, 173, 1602, 17575, 222497, 3188806, 50988405, \ldots$ (this is now sequence A75729 in [10]). Table I illustrates the case $n = 3$.

Several properties can be deduced from Theorem 1.

(i) By logarithmic differentiation of (1) (cf. [11, Chap. 1, p. 22]) we obtain a recurrence

$$H_n = \sum_{k=1}^{n} \binom{n-1}{k-1} B_k H_{n-k}\ . \tag{2}$$

(ii) Expanding the right-hand side of (1) leads to an explicit formula:

$$H_n = \sum_{(m_1, m_2, \ldots, m_n)} \frac{n!\ \prod_{j=1}^{n} B_j^{m_j}}{\prod_{j=1}^{n} m_j!\ (j!)^{m_j}}\ , \tag{3}$$

where the sum is over all $(m_1, m_2, \ldots, m_n)$ such that $\sum_{j=1}^{n} j m_j = n$; that is, m_j is the number of subsets of S containing j elements, for $j = 1, \ldots, n$.

When $n = 6$, for example,
$H_6 \;=\; \frac{6!}{6!(1!)^6}\ 1 + \frac{6!}{4!1!(1!)^4(2!)^1}\ 1^4\ 3^1 + \frac{6!}{2!2!(1!)^2(2!)^2}\ 1^2\ 3^2 + \frac{6!}{3!(2!)^3}\ 3^3 + \frac{6!}{3!1!(1!)^3(3!)^1}\ 1^3\ 13^1 + \frac{6!}{1!1!1!(1!)^1(2!)^1(3!)^1}\ 1^1\ 3^1\ 13^1 + \frac{6!}{2!1!(3!)^2}\ 13^2 + \frac{6!}{2!1!(1!)^2(4!)^1}\ 1^2\ 75^1 + \frac{6!}{1!1!(2!)^1(4!)^1}\ 3^1\ 75^1 + \frac{6!}{1!1!(1!)^1(5!)^1}\ 1^1\ 541^1 + \frac{6!}{1!(6!)^1}\ 4683^1 = 17575.$

(iii) The e.g.f. satisfies the differential equation

$$H'(x) \frac{(2 - \exp(x))^2}{\exp(x)} \;=\; H(x)\ .$$

(iv) Asymptotic behavior. The e.g.f. has an essential singularity at $x = \log 2$. The asymptotics are sufficiently complicated that we sought computer assistance. Salvy's Maple package *gdev* [5], [9] is specifically designed for this purpose. The result, after simplification, is that

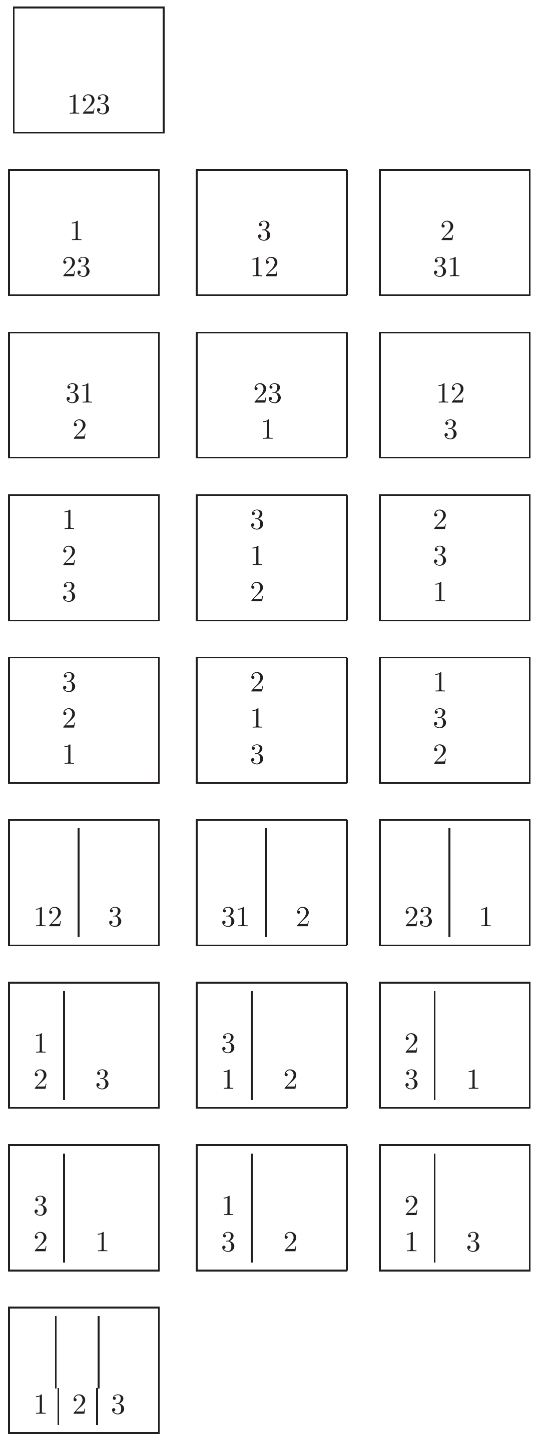
$$H_n \;\sim\; \frac{n!\, e^{\sqrt{2n/\log 2}}}{C^{1/4}\, n^{3/4}\, (\log 2)^n} \text{ as } n \to \infty\ , \tag{4}$$

structures in Section 2 are defined by the specification

$$[H, \{H = \mathit{Set}(\mathit{Sequence}(\mathit{Set}(Z, \mathit{card} >= 1), \mathit{card} >= 1))\}, \mathit{labeled}].$$

[10] now contains the sequences enumerating all such Motzkin-type structures in which the specification involves up to three occurrences of *Set* and *Sequence*.

Table I: For $n = 3$ elements there are 23 hierarchical orderings into at most three different subsets. The subsets are separated by bars, and the hierarchy within a subset is indicated by the vertical arrangement.



where

$$C \;=\; 32\,\pi^2\, \exp(3 - \frac{1}{\log 2})\, \log 2 \;=\; 1038.97\ldots\ .$$

This implies that

$$\log H_n \;\sim\; n \log n - n(1 + \log\log 2) + \sqrt{\frac{2n}{\log 2}} + O(\log n)\ .$$

For comparison, B_n satisfies

$$B_n \;\sim\; \frac{n!}{2\,(\log 2)^{n+1}}\ , \tag{5}$$

$$\log B_n \;\sim\; n \log n - n(1 + \log\log 2) + O(\log n)\ .$$

By iterating Theorem 1 we can also count hierarchies of hierarchical orderings. The e.g.f. is

$$\exp\left(\exp\left(\frac{1}{2 - \exp(x)} - 1\right) - 1\right)\ , \tag{6}$$

and the first few terms are $1, 1, 6, 52, 588, 8174, 134537, 2554647, 54909468, 1316675221, \ldots$ (sequence A75756 in [10]).

3 The unlabeled case

If the initial n elements are unlabeled, ordered set-partitions are called "compositions" of n, and their number is 2^{n-1} [8, p. 124]. To obtain a hierarchical ordering we partition the elements into $m \le n$ unlabeled and nonempty subsets, and form a composition of each subset. Let U_n denote the number of such hierarchical orderings, with ordinary generating function (or o.g.f.) $U(x) = \sum_{n \ge 0} U_n x^n$.

Theorem 2

$$U(x) \;=\; \prod_{j \ge 1} \frac{1}{(1 - x^j)^{2^{j-1}}}\ . \tag{7}$$

Proof. This is the unlabeled analogue of Theorem 1. If a_n, $n \ge 1$, is the number of n-element objects with a certain property, then b_n, the number of disjoint unions of such objects with a total of n elements, where the order of the components is unimportant, is given by

$$1 + \sum_{n=1}^{\infty} b_n x^n \;=\; \prod_{j=1}^{\infty} (1 - x^j)^{-a_j}$$

(see for example [3, p. 91]). ∎

The first few values U_n are $1, 1, 3, 7, 18, 42, 104, 244, 585, 1373, \ldots$ (sequence A34691 in [10]). When $n = 3$, for example, the $H_3 = 23$ hierarchical orderings in Table I reduce to $U_3 = 7$ when the labels are removed.

Properties. (i) Logarithmic differentiation of (7) leads to a recurrence:

$$U_n \;=\; \frac{1}{n}\sum_{k=1}^{n}\alpha_k U_{n-k}\,,\ \text{where } \alpha_k \;=\; \sum_{d|k} d2^{d-1}\;. \tag{8}$$

(ii) It is not so straightforward to find the asymptotic behavior in this case. This is to be expected, since the generating function for the number of partitions of n has a similar form to (7). Also, (7) does not belong to the family of generalized partition functions considered by Meinardus and discussed in [1, Chap. 7]. However, the saddle point method applies.

Theorem 3

$$U_n \;\sim\; \frac{2^n e^{\sqrt{2n}}}{\sqrt{2\pi}\,2^{3/4}\,e^{1/4}\,n^{3/4}}\ \textit{as } n\to\infty\;. \tag{9}$$

Proof. We have

$$\begin{aligned}\log U(x) \;&=\; \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{2^{n-1}\,x^{mn}}{m}\\ &=\; \sum_{N=1}^{\infty}x^N\sum_{d|N}\frac{d\,2^{d-1}}{N}\;,\end{aligned}$$

the interchange of summations being justified since all terms are positive, so

$$\log U(x) \;=\; \sum_{k=1}^{\infty}\frac{x^k}{1-2x^k}\;. \tag{10}$$

This has poles at $2^{-1}, 2^{-1/2}, 2^{-1/3}, \ldots$, and the radius of convergence is $1/2$. The pole at $x = 1/2$ dominates, and we apply the saddle-point method as in [7, §12]. The saddle point is at $x = r_n$, the solution to

$$r_n\,\frac{U'(r_n)}{U(r_n)} \;=\; n\;,$$

which is

$$r_n \;=\; \frac{1}{2}-\frac{1}{8n}\sqrt{8n+1}+\frac{1}{8n}+O(n^{-3/2})\;.$$

Then [7, Eq. (12.9)] leads to (9). ■

4 The structure of a random hierarchy

In this section we consider the case of a single (labeled or unlabeled) hierarchy. The elements at the bottom of the hierarchy will be said to have rank 1, those at the next level rank 2, and so on. The maximal rank in a hierarchy is its *height*.

Suppose there are n labeled elements. Let X be a hierarchy chosen at random from the B_n possibilities, and let $x \in X$ be a randomly chosen element. There are $h!\left\{ {n \atop h} \right\}$ ways that X can

have height h, where $\left\{ {n \atop h} \right\}$ is a Stirling number of the second kind (cf. [4, Chap. 7, Problem 44]), and indeed

$$B_n = \sum_{h=0}^{n} h! \left\{ {n \atop h} \right\} .$$

Given that X has height h, $x \in X$ is equally likely to have any rank from 1 to r. It follows that the probability that a randomly chosen x has rank r is

$$P_{n,r} = \sum_{i=r}^{n} \frac{1}{i} \frac{i! \left\{ {n \atop i} \right\}}{B_n} = \frac{1}{B_n} \sum_{i=r}^{n} (i-1)! \left\{ {n \atop i} \right\} . \tag{11}$$

The average rank is

$$a_n = \sum_{r=1}^{n} r P_{n,r} = \frac{1}{2B_n} \sum_{i=1}^{n} (i+1)! \left\{ {n \atop i} \right\} . \tag{12}$$

The numbers $a_n B_n$, $n = 0, 1, \ldots$, have e.g.f.

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{i=1}^{n} \frac{(i+1)!}{2} \left\{ {n \atop i} \right\} = -\frac{1}{2} \frac{(e^x-1)(e^x-3)}{(e^x-2)^2} ,$$

after some simplification, using standard properties of Stirling numbers (cf. [4, Table 351]). This implies

$$a_n B_n \sim n! \frac{n}{8(\log 2)^{n+2}} \text{ as } n \to \infty . \tag{13}$$

Combining this with (5) we find that the average rank is asymptotic to

$$\frac{n}{4 \log 2} = 0.36067 \ldots n \text{ as } n \to \infty . \tag{14}$$

In the unlabeled case, if one of the 2^{n-1} compositions of n is chosen at random, X (say), and $x \in X$ is picked at random, one can show that the probability that x has rank r is

$$\frac{1}{n\, 2^{n-1}} \sum_{i=r}^{n} \binom{n}{i} , \tag{15}$$

and the average rank is $(n+3)/4$. Ranks are higher in the labeled case.

Acknowledgements We thank Jeffrey Lagarias and Andrew Odlyzko for helpful discussions. T.W. thanks the Mathematics Department of the University Kassel for providing access to their computer pool.

References

[1] G. E. Andrews, *The Theory of Partitions*, Addison-Wesley, Reading, MA, 1976.

[2] F. Bergson, G. Labelle, P. Leroux, *Combinatorical Species and Tree-like Structures*, Encyclopedia of Mathematics and its Applications, Vol. **67**, Cambridge University Press, Cambridge, 1998.

[3] P. J. Cameron, Some sequences of integers, *Discrete Math.*, **75** (1989), 89–102.

[4] R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, Reading, MA, 1994, 2nd edition.

[5] INRIA Algorithms Project, *Algolib: The Algorithms Project's Library*, published electronically at pauillac.inria.fr/algo/libraries/, 2003.

[6] T. S. Motzkin, Sorting numbers for cylinders and other classification numbers, pp. 167–176 of T. S. Motzkin, ed., *Combinatorics*, Proc. Symp. Pure Math. **19**, Amer. Math. Soc., Providence, RI, 1971.

[7] A. M. Odlyzko, Asymptotic enumeration methods, pp. 1063–1229 of R. L. Graham, M. Groötschel and L. Lovász, eds., *Handbook of Combinatorics*, Elsevier, Amsterdam and MIT Press, Cambridge, MA, 1995, Vol. 2.

[8] J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, NY, 1958.

[9] B. Salvy, Examples of automatic asymptotic expansions, Report **114**, Institut National de Recherche en Informatique et en Automatique, 1989.

[10] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at www.research.att.com/∼njas/sequences/, 2003.

[11] H. S. Wilf, *Generatingfunctionology*, Academic Press, San Diego, 1994, 2nd edition.